JOBNAME: XJAS 62#2 2005 PAGE: 1 SESS: 62 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

Multiplicative Noise and Non-Gaussianity: A Paradigm for Atmospheric Regimes?

PHILIP SURA, MATTHEW NEWMAN, CÉCILE PENLAND, AND PRASHANT SARDESHMUKH

NOAA–CIRES Climate Diagnostics Center, Boulder, Colorado

(Manuscript received 27 January 2004, in final form 9 September 2004)

ABSTRACT

Atmospheric circulation statistics are not strictly Gaussian. Small bumps and other deviations from Gaussian probability distributions are often interpreted as implying the existence of distinct and persistent nonlinear circulation regimes associated with higher-than-average levels of predictability. In this paper it is shown that such deviations from Gaussianity can, however, also result from linear stochastically perturbed dynamics with multiplicative noise statistics. Such systems can be associated with much lower levels of predictability. Multiplicative noise is often identified with state-dependent variations of stochastic feedbacks from unresolved system components, and may be treated as stochastic perturbations of system parameters. It is shown that including such perturbations in the damping of large-scale linear Rossby waves can lead to deviations from Gaussianity very similar to those observed in the joint probability distribution of the first two principal components (PCs) of weekly averaged 750-hPa streamfunction data for the past 52 winters. A closer examination of the Fokker–Planck probability budget in the plane spanned by these two PCs shows that the observed deviations from Gaussianity can be modeled with multiplicative noise, and are unlikely the results of slow nonlinear interactions of the two PCs. It is concluded that the observed non-Gaussian probability distributions do not necessarily imply the existence of persistent nonlinear circulation regimes, and are consistent with those expected for a linear system perturbed by multiplicative noise.

1. Introduction

Atmospheric predictability is limited by the chaotic nature of weather. There are many different ways of defining predictability; for example, we may consider the doubling time of initial uncertainty. In forecasting, it is of considerable interest to find those initial atmospheric states for which this error doubling time is particularly long.

Attention has thus been paid to finding persistent or quasi-stationary atmospheric flow regimes, often loosely defined as large-scale atmospheric flow configurations that persist longer than individual weather systems (e.g., Pandolfo 1993). More specifically, regimes are identified with "preferred" regions of the atmospheric state space. The hope is that such regions contain initial atmospheric states from which better long-term forecasts are possible.

The idea of weather regimes, in terms of midlatitude cyclone tracks over Europe, was first proposed by van Bebber (1891). The more modern phenomenological notion of midlatitude atmospheric flow regimes, or *Großwetterlagen*, was introduced by Baur et al. (1944) and Baur (1947) for use in statistical long-range weather forecasting. Dynamical theories of multiple equilibria (blocked and zonal regimes) due to wave–mean flow interactions, first proposed by Charney and DeVore (1979), Wiin-Nielsen (1979), and Hart (1979), spurred a renewed effort to explore the existence of regimes in observed midlatitude flows (e.g., Sutera 1986; Hansen and Sutera 1986; Mo and Ghil 1988; Molteni et al. 1990; Kimoto and Ghil 1993a,b; Cheng and Wallace 1993; Corti et al. 1999; Smyth et al. 1999; Monahan et al. 2001; and many others). However, these observational studies generally did not find regime behavior as pronounced as in the theoretical and simple model studies. In fact, for monthly time scales, the observational evidence for regimes remains so weak that the null hypothesis of multinormal behavior cannot be rejected at the 5% confidence level (Stephenson et al. 2004).

The term "regime" can be nonspecific in common usage. It is often used merely to refer to a rapid transition to a quasi-stationary state. With this meaning, even some stable linear systems can exhibit "regime transitions" as a result of rapid singular-vector growth. Indeed, observed atmospheric regime transitions of this type can be well simulated by relatively low-dimensional linear models (Cash and Lee 2001; Winkler et al. 2001). To the extent that they are also good forecast models of weekly averages (Newman et al. 2003), such

Corresponding author address: Philip Sura, NOAA–CIRES Climate Diagnostics Center, R/CDC1, 325 Broadway, Boulder, CO 80305-3328.
E-Mail: Philip.Sura@noaa.gov

JOBNAME: XJAS 62#2 2005 PAGE: 2 SESS: 62 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

models imply an important constraint on the nature of nonlinearity in the atmosphere. We will return to this point later.

A stable linear system driven by Gaussian forcing will result in Gaussian statistics. Therefore, a necessary condition for regime behavior resulting from nonlinear dynamics is non-Gaussianity of the probability distribution function (PDF) representing all possible atmospheric states. As in the studies cited above, we will focus on this condition in this paper. Ideally, one would like to find significant multiple peaks in the full PDF. However, because of the limited data record, observational studies of climate regimes typically examine the bivariate PDF of the amplitudes of the two leading empirical orthogonal functions (EOFs) of an appropriate atmospheric variable (e.g., Mo and Ghil 1988; Molteni et al. 1990; Kimoto and Ghil 1993a,b; Corti et al. 1999; Smyth et al. 1999; Weisheimer et al. 2001). It is important to note again that such PDFs do not show any statistically significant multimodality. Rather, they show "inhomogeneities" (statistically significant deviations from bivariate Gaussianity) that are often interpreted as indicative of multiple Gaussian regimes (Smyth et al. 1999).

Models of regime behavior can be thought of as models of low-frequency atmospheric variability. They may be expressed as

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \mathbf{N_1}(\mathbf{x}, \mathbf{x}) + \mathbf{N_2}(\mathbf{x}, \mathbf{x}') + \mathbf{N_3}(\mathbf{x}', \mathbf{x}') + \mathbf{F}, \tag{1}$$

where $\mathbf{x}$ represents the resolved (e.g., low frequency or slow) portion of the atmospheric anomaly vector, and $\mathbf{x}'$ the unresolved (e.g., high frequency or fast) portion. Here, $\mathbf{Lx}$ denotes linear dynamics, and $\mathbf{N_1}$, $\mathbf{N_2}$, $\mathbf{N_3}$ nonlinear interactions between slow–slow, slow–fast, and fast–fast components, respectively; $\mathbf{F}$ represents external forcing. A common explanation for extratropical climate regimes is that they result from nonlinearities in the slow manifold of the equations governing atmospheric dynamics, that is, from $\mathbf{N_1}$ (e.g., Legras and Ghil 1985; Yoden 1985a,b; Ghil and Childress 1987; DeSwart 1988; Itoh and Kimoto 1996, 1997, 1999; Ghil and Robertson 2002; and many others). It is this slow process that gives hope for long-range predictability. Of course, slowly evolving external non-Gaussian forcing in $\mathbf{F}$, such as that produced by anomalous tropical convection due to the Madden–Julian oscillation (MJO) and El Niño–Southern Oscillation (ENSO), could also produce a non-Gaussian response in the extratropics even if the extratropical dynamics were linear.

Regimes that result from either or both of these mechanisms might indeed be more predictable than other atmospheric states. There is, however, a third possibility: regimes could result not from slow processes but rather from the fast (i.e., rapidly decorrelating) nonlinearities of the dynamical system contained in $\mathbf{N_2}$ and $\mathbf{N_3}$. Furthermore, given the very high number of degrees of freedom in the atmosphere and the consequent plethora of nonlinear subsystems, it is highly likely that there exists chaos sufficiently disordered as to make the application of the Central Limit Theorem (e.g., Khasminskii 1966; Papanicolaou and Kohler 1974; Sardeshmukh et al. 2001; Majda et al. 1999, 2003) valid at medium-range time scales as short as a week. In this case, the fast nonlinearities may be approximated as state-dependent, or multiplicative, stochastic noise that is inherently unpredictable [for state-independent, or additive, noise this approximation was first noted in a climate context by Hasselmann (1976)].

That apparently similar regimes can result from either slow or fast nonlinearities is illustrated in Fig. 1 (a more quantitative discussion is given in section 3). Consider a double potential well as a simple model of two regimes resulting from nonlinearities of the slow manifold. System transitions from one potential well to the other by additive noise kicks will result in a bimodal PDF. This is not, however, the only dynamical system that can produce such a PDF. Consider instead a linear system, represented by a unimodal deterministic potential, in which the state trajectories are perturbed by multiplicative noise. If the noise is relatively stronger in the center of the potential well than at the edges, then this system will also have a bimodal PDF. Thus the same bimodal PDF can result from either a slow (deterministic) nonlinear dynamical system or a fast (stochastic) nonlinear dynamical system. As we will see, however, the predictability of these two systems is very different.

In this paper we use both simple models and observational analysis to develop the hypothesis that atmospheric non-Gaussian regimes may be due to multiplicative noise. First, in section 2, some principal results of stochastic dynamics are briefly reviewed. Then, in section 3, we discuss some simple examples that demonstrate how a given PDF can imply very different predictability depending upon whether the PDF results from slow deterministic nonlinear evolution or fast stochastic evolution. In section 4, stochastic perturbations of the linear damping coefficient and the ambient zonal mean flow are considered in a linear model of Rossby wave evolution to assess the impact of multiplicative noise in a simple, but meteorologically relevant, setting. By explicitly solving the Fokker–Planck equation and the stochastic differential equation, we show that multiplicative noise in the frictional damping leads to intermittency and consequently a highly non-Gaussian distribution. These results provide one possible explanation for the non-Gaussian PDF found in the observational analysis presented in section 5, where it is shown that the regime behavior in the leading EOFs of 750-hPa Northern Hemisphere streamfunction is consistent with that resulting from multiplicative noise. Finally, section 6 provides a summary and a discussion.

JOBNAME: XJAS 62#2 2005 PAGE: 3 SESS: 62 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

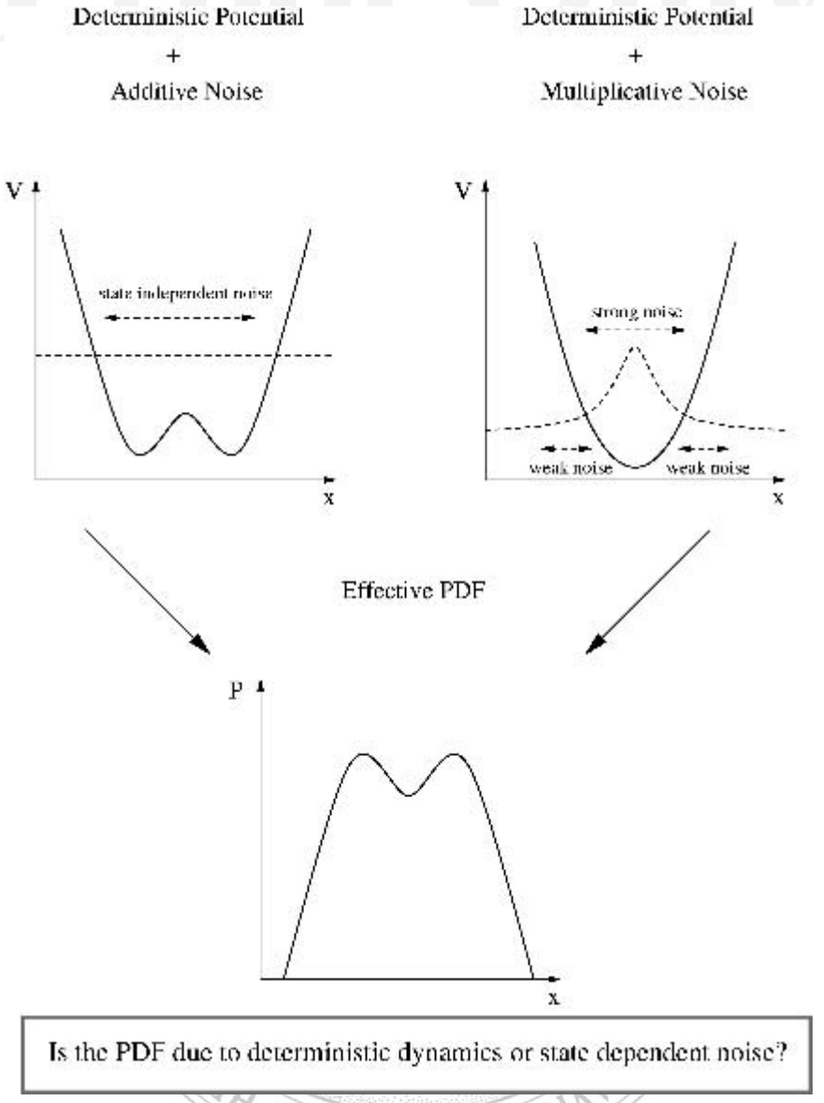

FIG. 1. A schematic sketch to illustrate the fundamental dynamical difference between deterministically and stochastically induced regimes. The effective PDF of a trajectory in a deterministic double-well potential driven by additive noise will be bimodal. The same effective PDF can be produced by a trajectory in a unimodal deterministic potential kicked around by multiplicative noise. Because of the larger noise amplitudes near the center of the monomodal potential, as compared to the strength of the noise right and left of it, the system's trajectory is more often found on either side of the central noise maximum. Thus, the PDF becomes bimodal. See appendix A for a mathematical formulation of this behavior.

2. Stochastic dynamics in a nutshell

This section reviews a few basic ideas of stochastic dynamics used in this paper. More comprehensive treatments may be found in many textbooks (e.g., Gardiner 1985; Horsthemke and Léfèver 1984; Paul and Baschnagel 1999).

Consider the dynamics of an n-dimensional system whose state vector $\mathbf{x}$ is governed by the stochastic differential equation (SDE)

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\boldsymbol{\eta}, \tag{2}$$

where the vector $\mathbf{A}$ represents all slow processes and $\mathbf{B}(\mathbf{x})\eta$, with the matrix $\mathbf{B}$ and the noise vector η, represents the stochastic approximation to the fast nonlinear processes. The stochastic components η_i are assumed to be independent Gaussian white noise processes:

JOBNAME: XJAS 62#2 2005 PAGE: 4 SESS: 62 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

$$\langle \eta_i(t) \rangle = 0, \quad \langle \eta_i(t)\eta_i(t') \rangle = \delta(t - t'), \tag{3}$$

where $\langle \ldots \rangle$ denotes the averaging operator. The corresponding Fokker–Planck equation,

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}\left[A_i + \alpha \sum_{j,k} \left(\frac{\partial}{\partial x_j} B_{ik} \right) B_{jk} \right] p(\mathbf{x}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\mathbf{B}\mathbf{B}^{\mathrm{T}})_{ij}\, p(\mathbf{x}, t), \tag{4}$$

describes the conservation of the probability density $p(\mathbf{x}, t)$ of the system described by the SDE (2). Two different values of α yield two physically important stochastic calculi: the Itô ($\alpha = 0$) and the Stratonovich calculus ($\alpha = 1/2$). On the right-hand side, the first term within square brackets describes the dynamics of the deterministic system and is called the deterministic drift. The second term within square brackets, which does not occur in Itô systems ($\alpha = 0$), is called the noise-induced drift. The remaining term is associated with the diffusion of the probability density by noise.

For a detailed discussion of stochastic integration and the differences between Itô and Stratonovich SDEs, see Horsthemke and Léfèver (1984), Gardiner (1985), or Penland (1996). The key point here is that the Stratonovich calculus is relevant for continuous physical systems, such as the atmosphere, in which rapidly fluctuating quantities with small but finite correlation times are approximated as white noise. Thus, simplified stochastic models constructed from atmospheric dynamical equations may assume Stratonovich calculus. However, if a stochastic model is indirectly estimated from observed discrete data, then the inferred drift will be the sum of the deterministic and the noise-induced drifts. In this case, using the Itô framework may be preferable, where now $\mathbf{A}(\mathbf{x})$ represents not just the deterministic drift but rather this sum, or the "effective drift."

Equations for moments of $\mathbf{x}$ can be obtained by multiplying the Fokker–Planck Eq. (4) by powers of $\mathbf{x}$ and integrating over all $\mathbf{x}$. In particular, second moments of $\mathbf{x}$ are given by

$$\frac{d\langle \mathbf{x}\mathbf{x}^{\mathrm{T}} \rangle}{dt} = \langle \mathbf{A}(\mathbf{x})\mathbf{x}^{\mathrm{T}} \rangle + \langle \mathbf{x}\mathbf{A}^{\mathrm{T}}(\mathbf{x}) \rangle + \langle \mathbf{B}(\mathbf{x})\mathbf{B}^{\mathrm{T}}(\mathbf{x}) \rangle. \tag{5}$$

This equation is known as the fluctuation–dissipation relation (FDR) of the system (see e.g., Penland and Matrosova 1994).

In principle, the deterministic and stochastic parts of (4) can be determined from data by using their statistical definitions (Siegert et al. 1998; Friedrich et al. 2000; Gradišek et al. 2000; Sura and Barsugli 2002; Sura 2003; Sura and Gille 2003):

$$\mathbf{A}(\mathbf{x}) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \langle \mathbf{X}(t + \Delta t) - \mathbf{x} \rangle |_{\mathbf{X}(t)=\mathbf{x}} \tag{6}$$

$$\mathbf{B}(\mathbf{x})\mathbf{B}^{T}(\mathbf{x}) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \langle [\mathbf{X}(t + \Delta t) - \mathbf{x}][\mathbf{X}(t + \Delta t) - \mathbf{x}]^{\mathrm{T}} \rangle |_{\mathbf{X}(t)=\mathbf{x}}, \tag{7}$$

where $\mathbf{X}(t + \Delta t)$ is a solution (a single stochastic realization) of the SDE (2) with the initial condition $\mathbf{X}(t) = \mathbf{x}$ at time t. The data define a state space representing every observed value of $\mathbf{x}$. The effective drift and stochastic diffusion are estimated by replacing the theoretical limit $\Delta t \to 0$ with a finite-difference approximation. In practice, estimating $\mathbf{B}(\mathbf{x})\mathbf{B}^{\mathrm{T}}(\mathbf{x})$ from discretely sampled data is prone to error, because Taylor expansions of stochastic terms are proportional to $\sqrt{\Delta t}$ and not proportional to Δt as are the deterministic terms (e.g., Sura and Barsugli 2002; Sura 2003). Note that $\mathbf{B}(\mathbf{x})\mathbf{B}^{\mathrm{T}}(\mathbf{x})$ rather than $\mathbf{B}(\mathbf{x})$ is estimated from data. In general, it is impossible to find a unique expression for $\mathbf{B}(\mathbf{x})$ in the multivariate case (e.g., Monahan 2004).

When $\mathbf{A}$ and $\mathbf{B}$ are known, analytical solutions of the Fokker–Planck Eq. (4) for $p(\mathbf{x}, t)$ can only be found in limited cases (appendix B presents one such case). For more general cases, numerical methods must be used. To interpret the results of the Fokker–Planck equation, numerical integrations of the SDE (2) can also be performed.

3. Two paradigms for atmospheric regimes

Obviously, a given non-Gaussian PDF can represent many different dynamical systems. However, it is illustrative to consider two extreme models: a deterministic model, in which regimes are entirely due to a nonlinear deterministic $\mathbf{A}$ (perturbed only by state-independent noise), and a stochastic model, in which regimes are entirely due to a multiplicative noise term $\mathbf{B}$ (with only a linear $\mathbf{A}$). That two such models can produce the same PDF was highlighted in Fig. 1. The mathematical details of the two models are given in appendix A.

First, we quantify the simple example presented in Fig. 1 by applying these models to the bimodal PDF $p(x) = (1/\sqrt{8\pi})\{\exp[-(x + 1.5)^2/2] + \exp[-(x - 1.5)^2/2]\}$ (see Fig. 2a). Because observed PDFs rarely show any clear multimodality, we also consider the "skewed" PDF given in Fig. 3a, whose departure from Gaussianity is relatively small but isconsistent with that observed for weekly averaged circulation anomalies (see section 5). It has a heavier tail than a Gaussian for values $x < -2$ (regime 1), is smaller than a Gaussian for $-2 < x < -0.3$, and is again heavier than a Gaussian for $-0.3 < x < 0.5$ (regime 2). For larger x the PDF is strictly Gaussian.

Given the PDF and $B^2 \equiv 1$, we can solve for $A(x)$ [Eq. (A.3); see appendix A]; results for the bimodal PDF are in Fig. 2b and for the skewed PDF in Fig. 3b. In this case the non-Gaussianity is due to the nonlinearity of the deterministic term $A(x)$. Alternatively, given the PDF and $A(x) = -\varepsilon x$, we can solve for $B(x)$

JOBNAME: XJAS 62#2 2005 PAGE: 5 SESS: 62 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

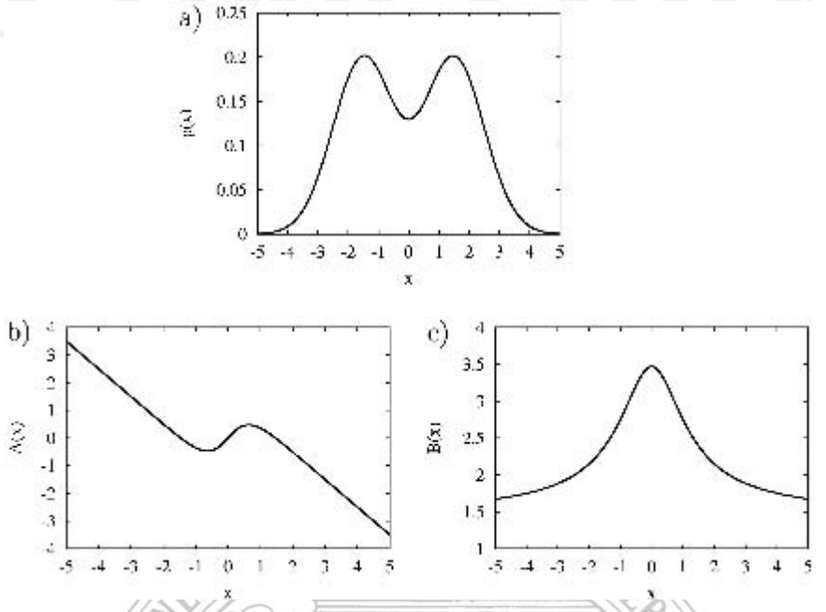

FIG. 2. (a) Graph of the bimodal PDF $p(x) = (1/\sqrt{8\pi})[\exp(-(x + 1.5)^2/2) + \exp(-(x - 1.5)^2/2)]$. (b) Solution of (A.3), given the bimodal PDF and additive noise $B^2 \equiv 1$. (c) Solution of (A.4), given the bimodal PDF and a linear deterministic term $A(x) = -x$.

[Eq. (A.4)]; results for the bimodal PDF for $\varepsilon = 1$ are in Fig. 2c and for the skewed PDF in Fig. 3c. Now the non-Gaussianity is due to the structure of the multiplicative noise term $B(x)$. Note (see Fig. 3c) that $B(x)$ for the skewed PDF is approximately piecewise linear, and that as opposed to the stochastic model for the bimodal PDF, the noise amplitude increases for decreasing negative x.

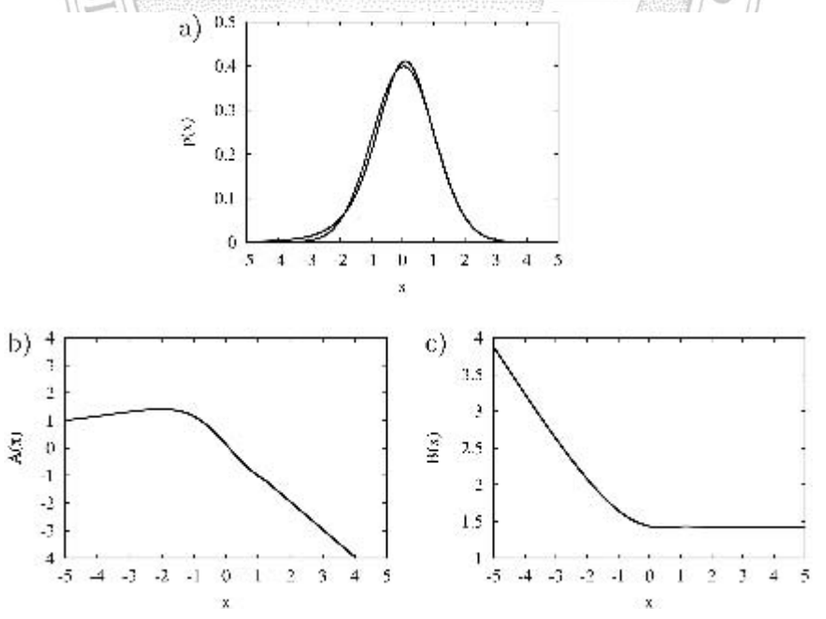

FIG. 3. (a) Graph of the non-Gaussian PDF (solid line); the corresponding Gaussian PDF is indicated by the dashed line. (b) Solution of (A.3), given the PDF and additive noise $B^2 \equiv 1$. (c) Solution of (A.4), given the PDF and a linear deterministic term $A(x) = -x$.

JOBNAME: XJAS 62#2 2005 PAGE: 6 SESS: 69 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

Although the stationary PDFs of the deterministic and stochastic regimes are identical, the conditional PDFs (the probability of an event given that another event has occurred) and related mean residence times are not. Here, we define residence time as the time it takes a stochastic trajectory initially at x in the interval $[x_1, x_2]$ to first leave that interval (see appendix A for more details). For example, the mean residence time in the right peak of the bimodal PDF (interval [1, 2]), shown in Fig. 4a, is considerably longer for the deterministic model than for the stochastic model. A similar difference between the two models exists in the interval [−3.5, −2.5] for the weakly skewed PDF (Fig. 5a), but is much less for the interval [−0.25, 0.75] (Fig. 5b).

Not surprisingly, the predictability in these two systems is also very different. Here, we define predictability by the expected skill of a perfect model infinite-member forecast ensemble, measured as an anomaly correlation:

$$\rho_\infty(\tau) = \frac{S(\tau)}{\sqrt{S(\tau)^2 + 1}}, \tag{8}$$

where $S(\tau) \equiv s(\tau)/\sigma(\tau)$ is the signal-to-noise ratio and τ is the forecast lead (for example, Sardeshmukh et al. 2000; Newman et al. 2003). Here, the signal $s(\tau)$ is the ensemble mean, and the noise $\sigma(\tau)$ is the ensemble standard deviation, and $\rho_\infty(\tau)$ is the expected skill of a perfect model in which the signal is determined as the mean of an infinite-member ensemble [see Sardeshmukh et al. (2000) or Newman et al. (2003) for a more detailed discussion]. For the bimodal PDF, $\rho_\infty(\tau)$ for the initial condition $x_0 = 1.5$ for the deterministic and the stochastic models, is shown in Fig. 4b; $\rho_\infty(\tau = 2)$ as a function of initial condition is shown in Fig. 4c. As implied by the mean residence times, predictability is much less for the bimodality than that due to deterministic nonlinear dynamics because of unpredictable multiplicative noise. For the skewed PDF, $\rho_\infty(\tau)$ for the initial conditions $x_0 = -3$, and $x_0 = 0.25$ are shown in Figs. 5c,d; $\rho_\infty(\tau = 2)$ as a function of initial condition is shown in Fig. 5e. Note in Fig. 5d that $\rho_\infty(\tau)$ for the deterministic and stochastic models (and the initial condition $x_0 = 0.25$) are almost identical. Interestingly, predictability for the deterministic model is generally higher for negative x, whereas it is generally higher for positive x in the stochastic model. Thus, a characterization of predictability for a system represented by a given marginal (that is, unconditional) PDF depends on the dynamics of the underlying system, and cannot be inferred merely from the non-Gaussianity of the marginal PDF.

One point remains to be discussed. It seems possible that the lower predictability of the stochastic models might be entirely due to a shorter autocorrelation time scale because we arbitrarily prescribed a certain deterministic time scale by using $A(x) = -\varepsilon x$ with $\varepsilon = 1$ in Eq. (A.4). For the bimodal example, setting $\varepsilon = 0.15$

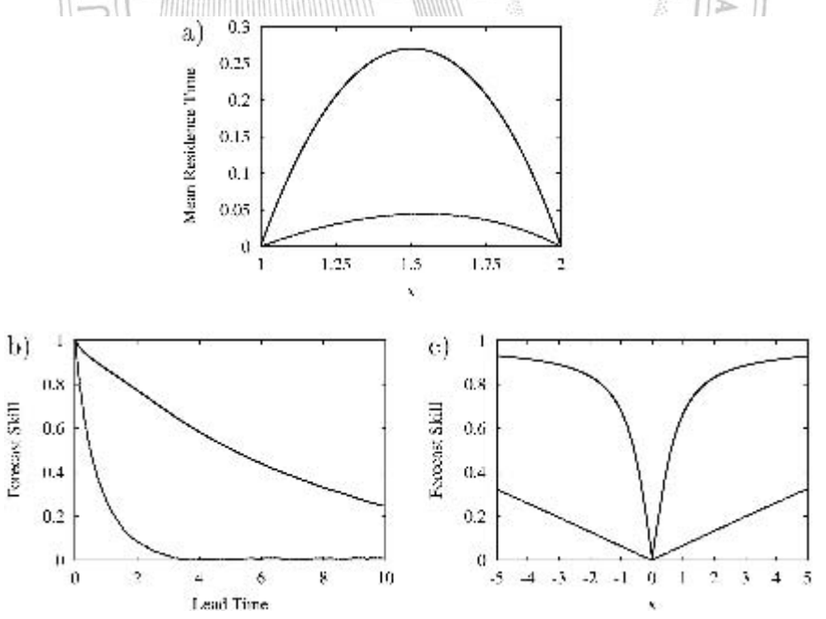

FIG. 4. Statistics of the bimodal case (see Fig. 2). (a) The mean residence time for the deterministic (sold line) and the stochastic model (dashed line) in the interval [1, 2], (b) ρ_∞ as a function of lead time for the deterministic (solid line) and the stochastic model (dashed line) for the initial condition $x_0 = 1.5$, and (c) ρ_∞ as a function of initial condition for the deterministic (solid line) and the stochastic model (dashed line) for the lead time $\tau = 2$.

JOBNAME: XJAS 62#2 2005 PAGE: 7 SESS: 62 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

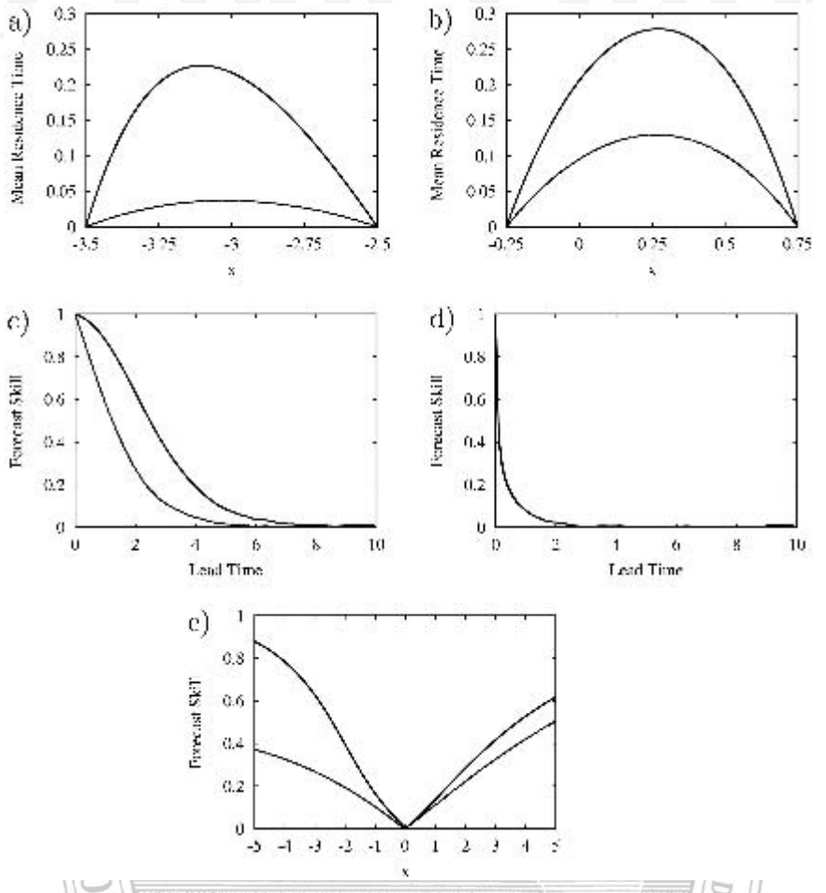

FIG. 5. Statistics of the non-Gaussian case (see Fig. 3). The mean residence times for the deterministic (solid lines) and the stochastic model (dashed lines) in the intervals (a) $[-3.5, -2.5]$, and (b) $[-0.25, 0.75]$. Also ρ_∞ as a function of lead time for the deterministic (solid line) and the stochastic model (dashed line) for the initial conditions (c) $x_0 = -3$, and (d) $x_0 = 0.25$ (note that $\rho_\infty(\tau)$ for the deterministic and stochastic model are almost identical), and (e) ρ_∞ as a function of initial condition for the deterministic (solid line) and the stochastic model (dashed line) for the lead time $\tau = 2$.

ensures that the multiplicative noise model has the same autocorrelation function as the deterministic model. But, even then, the stochastic model is less predictable than the deterministic model. For the more realistic skewed example, however, the stochastic and the deterministic models have virtually identical autocorrelation functions. That is, in both the bimodal and the skewed examples, the difference in predictability is also due to the structure of the noise and not only the damping time scale of the deterministic term.

4. Stochastically perturbed Rossby waves

In stochastic atmospheric models, noise is introduced primarily as an additive process (e.g., Egger 1981; DeSwart 1988; DelSole and Farrell 1995; Farrell and Ioannou 1996; Newman et al. 1997; Whitaker and Sardeshmukh 1998). However, stochastic forcing may, for example, also represent the fluctuations of model parameters due to unresolved system components (e.g., Neelin and Weng 1999; Sardeshmukh et al. 2001; Sura 2002; Lin and Neelin 2002; Monahan 2004). In that case, the stochastic process appears as multiplicative noise which, as is well known, can substantially change the dynamical behavior of not only nonlinear systems (Horsthemke and Léfèver 1984; Landa and McClintock 2000; Sura 2002), but also linear systems. The potentially significant role of multiplicative noise to improve the representation of subgrid-scale phenomena in models of the climate systems has been stressed in several studies (e.g., Buizza et al. 1999; Palmer 2001; Sardeshmukh et al. 2001; Pérez-Muñuzuri et al. 2003; Sura 2002). Sardeshmukh et al. (2001) introduced multiplicative noise in the linearized barotropic vorticity equa-

JOBNAME: XJAS 62#2 2005 PAGE: 8 SESS: 63 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

tion, and found that the mean stationary wave response to steady forcing was amplified when the damping parameter fluctuated, but was weakened (in a scale-dependent manner) when the ambient flow fluctuated. Here, we show in the same framework that stochastic damping also results in Rossby waves with a highly non-Gaussian distribution.

4.a. *Multiplicative noise in the linearized barotropic vorticity equation*

The linearized barotropic vorticity equation is:

$$\frac{\partial \zeta'}{\partial t} = -\nabla \cdot (\bar{\mathbf{v}}\zeta' + \mathbf{v}'\bar{\zeta}) - r\zeta' + F', \tag{9}$$

where ζ is absolute vorticity, $\mathbf{v}$ is the nondivergent horizontal velocity, r is the frictional damping rate, and F' is anomalous forcing; F' includes both predictable forcing (for example, steady tropical forcing) and unpredictable forcing modeled using additive noise (Newman et al. 1997). Overbars indicate time means, and primes denote deviations from the time means.

Equation (9) can be written in terms of the streamfunction $\boldsymbol{\Psi}$ ($\zeta' = \nabla^2\boldsymbol{\Psi}'$) and in operator form as

$$\frac{d\boldsymbol{\Psi}}{dt} = \mathbf{L}\boldsymbol{\Psi} + \mathbf{F}, \tag{10}$$

where $\mathbf{L}$ is the linear barotropic operator (for example, Borges and Sardeshmukh 1995) and (dropping the primes) $\boldsymbol{\Psi}$ and $\mathbf{F}$ are the anomalous streamfunction and forcing vectors, respectively. Any streamfunction (or forcing) anomaly can then be expanded in the basis set of complex eigenvectors $\mathbf{E}_j$ of $\mathbf{L}$ as $\boldsymbol{\Psi} = \Sigma\Psi_j\mathbf{E}_j$, where $\mathbf{L}\mathbf{E}_j = \lambda_j\mathbf{E}_j$, λ_j are the complex eigenvalues, and Ψ_j are the complex expansion coefficients. The equations for any mode are decoupled from those of all other modes so that for each mode j we can write

$$\frac{d\boldsymbol{\Psi}_\mathbf{j}}{dt} = \mathbf{L}_\mathbf{j}\boldsymbol{\Psi}_\mathbf{j} + \mathbf{f}_\mathbf{j}, \tag{11}$$

where now the vectors $\boldsymbol{\Psi}_j$ and $\mathbf{f}_\mathrm{j}$ represent $(\Psi_{jr}, \Psi_{ji})^\mathrm{T}$ and $(f_{jr}, f_{ji})^\mathrm{T}$ respectively, and

$$\mathbf{L}_\mathbf{j} = \begin{pmatrix} \lambda_{jr} & -\lambda_{ji} \\ \lambda_{ji} & \lambda_{jr} \end{pmatrix}. \tag{12}$$

For simplicity we now omit the index j.

Because this model does not exhibit variability beyond the decay of damped waves to a steady state, it is assumed that the model can be made more realistic by introducing stochastic perturbations. On the one hand, the modes retained in the model can be perturbed by additive noise. On the other hand, introducing stochastic perturbations in either λ_r or λ_i results in multiplicative noise, and from (11) the Stratonovich SDE is

$$\frac{d\boldsymbol{\Psi}}{dt} = \mathbf{L}\boldsymbol{\Psi} + \mathbf{f} + \mathbf{B}\boldsymbol{\eta}^M + \boldsymbol{\eta}^A, \tag{13}$$

where the forcing is separated into a deterministic component $\mathbf{f}$ and an additive stochastic component η^A. If the decay rate of the wave (λ_r) is perturbed stochastically, then

$$\mathbf{B} = \begin{pmatrix} -\Psi_r & 0 \\ -\Psi_i & 0 \end{pmatrix}. \tag{14}$$

If, instead, the phase speed (λ_i) is perturbed stochastically, then $\mathbf{B}$ is

$$\mathbf{B} = \begin{pmatrix} -\Psi_i & 0 \\ \Psi_r & 0 \end{pmatrix}. \tag{15}$$

In either case, the multiplicative (M) and additive (A) stochastic noise vectors are $\boldsymbol{\eta}^M = (\eta_r^M, \eta_i^M)^\mathrm{T}$, and $\boldsymbol{\eta}^A = (\eta_r^A, \eta_i^A)^\mathrm{T}$. In the subsequent discussion $\eta_r^M, \eta_i^M, \eta_r^A$, and η_i^A are assumed to be independent Gaussian white-noise processes with corresponding amplitudes $\sigma_r^M, \sigma_i^M, \sigma_r^A$, and σ_i^A:

$$\begin{aligned} &\langle \eta_r^M(t)\rangle = 0, \ \langle \eta_r^M(t)\eta_r^M(t')\rangle = (\sigma_r^M)^2\delta(t-t') \\ &\langle \eta_i^M(t)\rangle = 0, \ \langle \eta_i^M(t)\eta_i^M(t')\rangle = (\sigma_i^M)^2\delta(t-t') \\ &\langle \eta_r^A(t)\rangle = 0, \ \langle \eta_r^A(t)\eta_r^A(t')\rangle = (\sigma_r^A)^2\delta(t-t') \\ &\langle \eta_i^A(t)\rangle = 0, \ \langle \eta_i^A(t)\eta_i^A(t')\rangle = (\sigma_i^A)^2\delta(t-t') \end{aligned}, \tag{16}$$

where $\langle \ldots \rangle$ denotes the averaging operator. The PDF of this stochastic model is governed by the corresponding Fokker–Planck Eq. (4) with $\alpha = 1/2$.

4.b. *Stochastically perturbed Rossby waves on a superrotating basic state*

The simplest form of Eq. (9) sets the time mean flow to be in solid body rotation with the earth, where the mean meridional flow is zero and the mean zonal flow is $\bar{u} = u_0\cos(\theta)$, where u_0 is constant and θ is latitude. Then the eigenmodes $\mathbf{E}_j$ are just complex spherical harmonics $Y_n^m(\theta,\phi)$ for the wavenumber index pair (m, n) where φ is longitude, m is zonal wavenumber, and n is meridional wavenumber. The Rossby wave dispersion relation gives $\lambda_r = -r$ and $\lambda_i = mD_n$, where

$$\begin{aligned} D_n &= \left[\frac{2}{n(n+1)} + d_n\Delta\right]\Omega \\ d_n &= \frac{2}{n(n+1)} - 1 \end{aligned}, \tag{17}$$

(see Sardeshmukh et al. 2001). In the example below, typical parameters of atmospheric motion are used: $r = (4\ \mathrm{day})^{-1}$, $u_0 = 15\ \mathrm{m\ s^{-1}}$, $\Delta = u_0/a\Omega = 0.0323$, $\Omega = 2\pi\ \mathrm{day}^{-1}$, and mD_n is set to $1\ \mathrm{day}^{-1}$, corresponding to $m = O(5)$ and $n = O(5)$. Steady anomalous forcing is $\mathbf{F} = (1, 1)^\mathrm{T}$.

Past studies examining the Rossby wave response to steady forcing (Hoskins et al. 1977; Sardeshmukh and Hoskins 1988) have typically considered r and u_0 to be fixed. However, a more realistic representation of Rossby wave propagation on the sphere might also con-

JOBNAME: XJAS 62#2 2005 PAGE: 9 SESS: 64 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

sider stochastic perturbations in r and u_0 (Sardeshmukh et al. 2001). In the following, we solve the Fokker–Planck Eq. (4) for p using these parameters and stochastically perturbing either r or u_0 for different values of the noise amplitude σ_r^M. Weak additive noise with amplitudes $\sigma_r^A = \sigma_i^A = 0.2$ is also included. Note that because of the structure of the matrices (14) and (15), the imaginary part of the multiplicative noise η_i^M (with amplitude σ_i^M) has no impact and is, therefore, not specified. In both experiments the strength of the multiplicative noise σ_r^M is scaled by its effect on the variance of the system. The multiplicative noise amplitude increased until the variances $\langle \Psi_r^2 \rangle$ and $\langle \Psi_i^2 \rangle$ reach approximately 1. This yields the interval $\sigma_r^M = 0.0,0.1,\ldots 0.4$ for the stochastic damping, and the interval $\sigma_r^M = 0.0,0.2,\ldots 0.8$ for the stochastic basic state.

The semi-implicit Chang–Cooper method is implemented to solve (4) for $p(\mathbf{x}, t)$ (Chang and Cooper 1970; Park and Petrosian 1996). It employs a flux-conservative second-order accurate finite difference scheme, extended to multidimensional problems using the operator splitting method (Park and Petrosian 1996; Press et al. 1992). A regular grid with a mesh size 0.1 and 200 $\times$ 200 grid points is used. The domain of computations is [−10: 10, −10: 10]. The Fokker–Planck equation is integrated until a steady state is reached. The initial distribution of the probability density function p is chosen to be a two-dimensional Gaussian, with standard deviation $\sqrt{0.1}$, centered at the origin of the system (the results are not sensitive to this choice). The value of p integrated over the domain of computation is conserved and normalized to 1. To interpret the results of the Fokker–Planck equation, numerical integrations of the SDE are also performed, using the Milstein scheme (Kloeden and Platen 1992) with a time step of 0.1 days. Choosing an accurate scheme is especially important for the numerical integration of an SDE with multiplicative noise (Ewald et al. 2004). PDFs estimated from numerical integrations of the SDE agreed very well with solutions of the Fokker–Planck equation. Although a single experiment is only one realization of a stochastic process, the ergodic nature of the system ensures agreement with the steady state solution of the corresponding Fokker–Planck equation for the PDF of an infinite number of stochastic realizations.

4.B.1. STOCHASTIC DAMPING

The marginal PDFs for varying noise amplitude in the damping are shown in Figs. 6a,b. Table 1 shows the values of the skewness and the excess kurtosis (i.e., the excess over the kurtosis value of 3 for a Gaussian distribution). If the multiplicative noise is nonexistent or weak ($\sigma_r^M = 0.0, 0.1, 0.2$), the marginal distributions $p(\Psi_r)$ and $p(\Psi_i)$ are approximately Gaussian. However, for stronger noise amplitudes ($\sigma_r^M = 0.3, 0.4$) the distributions become skewed with heavy, highly non-Gaussian tails. A sample time series of Ψ_i of the stochastic Rossby wave model with relatively strong stochastic frictional damping ($\sigma_r^M = 0.4$) is shown in Fig. 7a. The distinct feature of the time series is its inter-

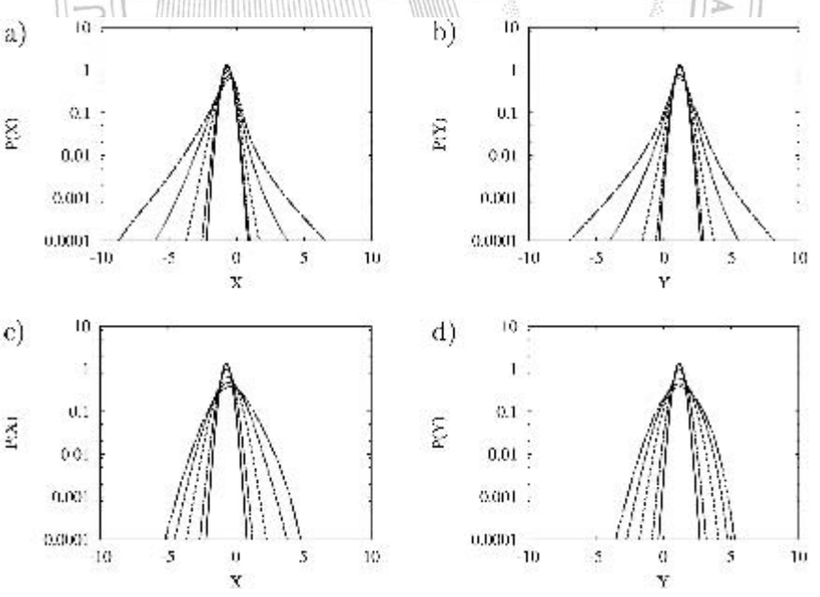

FIG. 6. (a),(b) Steady state marginal PDF of the Rossby wave model with stochastic frictional damping for $\sigma_r^M = 0.0$ (solid line), 0.1 (long dashed line), 0.2 (short dashed line), 0.3 (dotted line), and 0.4 (dotted–dashed line). (a) $p(\Psi_r)$, X denotes Ψ_r, (b) $p(\Psi_i)$, Y denotes Ψ_i. (c),(d) Steady state marginal PDF of the Rossby wave model with stochastic basic state for $\sigma_r^M = 0.0$ (solid line), 0.2 (long dashed line), 0.4 (short dashed line), 0.6 (dotted line), and 0.8 (dotted–dashed line), (c) $p(\Psi_r)$, X denotes Ψ_r, (d) $p(\Psi_i)$, Y denotes Ψ_i.

JOBNAME: XJAS 62#2 2005 PAGE: 10 SESS: 67 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

TABLE 1. Skewness and excess kurtosis of the marginal distributions $p(\Psi_r)$ and $p(\Psi_i)$ of the stochastic Rossby wave model with stochastic frictional damping. The excess kurtosis measures the departure from Gaussianity: excess kurtosis = kurtosis − 3. Note the strong departures (intermittency) from Gaussianity.

	Skewness		Excess kurtosis	
Multiplicative noise σ_r^M	Ψ_r	Ψ_i	Ψ_r	Ψ_I
0.0	0.0	0.0	0.0	0.0
0.1	−0.1	0.0	0.2	0.2
0.2	−0.3	−0.1	1.4	1.3
0.3	−0.7	−0.3	6.6	7.5
0.4	−2.7	−1.4	99.7	83.7

mittent behavior, which gives rise to the heavy, non-Gaussian tails of the PDFs, meaning that extreme events are far more probable than can be expected from a Gaussian distribution. The positive excess kurtosis is a measure of this intermittency.

4.b.2. STOCHASTIC BASIC STATE

Unlike the stochastic damping case, increasing noise in the superrotation does not produce heavy-tailed distributions, as is evident in both the marginal PDFs (Figs. 6c,d) and the skewness and excess kurtosis in Table 2. A sample time series of Ψ_i with a relatively strong stochastic basic state ($\sigma_r^M = 0.8$) (Fig. 7b) does not show the intermittent behavior seen in Fig. 7a. Because the stochasticity in the basic state only influences the phase of the Rossby wave, the PDF of the system remains more nearly Gaussian, while multiplicative noise in the frictional damping alters the energy of the wave and causes the PDF to be highly non-Gaussian (see appendix B and Fig. B.1 for more details).

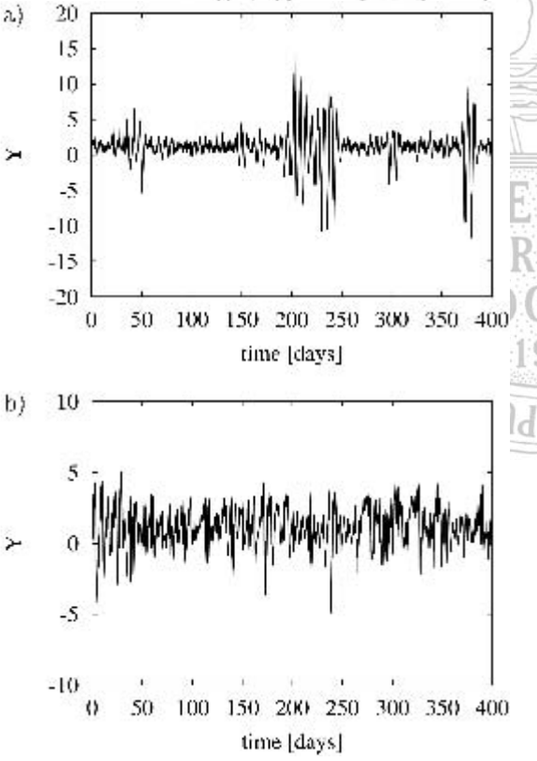

FIG. 7. Sample time series of Ψ_i of the stochastic Rossby wave model with (a) stochastic frictional damping ($\sigma_r^M = 0.4$) and (b) stochastic basic state ($\sigma_r^M = 0.8$). Here, Y denotes Ψ_i. Note the different scales on the ordinates.

4.c. *Departures from bivariate Gaussianity*

As can be expected from the marginal PDFs, stochastic damping produces notable departures from bivariate Gaussianity. This is true not only for parameters representing Rossby waves on a superrotating flow, but also for parameters representing the more general problem of large-scale barotropic Rossby waves evolving on a zonally and meridionally varying base state. As an example, we set $\lambda_r = 0.07$ and $\lambda_i = 0.19$ in (12), corresponding to a period of about 33 days and an *e*-folding time of 14 days, typical of the least-damped eigenmode of the 250-hPa climatological December–February flow (Borges and Sardeshmukh 1995). We set $\sigma_r^M = 0.05$, and for simplicity, set the additive noise to zero; including additive noise tends to decrease skewness, which can be offset by increasing **F**, but has much less effect upon the heavy tails. For display purposes, we remove the mean drift, rotate the PDF so that the abscissa points in the direction of maximum covariance, normalize by the variance along both axes, and show the difference between the PDF and the corresponding bivariate Gaussian distribution.

The resulting departure from Gaussianity (Fig. 8) appears as two slightly arcing ridge/trough (positive/negative departures) pairs aligned roughly along a line. This pattern is qualitatively similar for a wide parameter range. Decreasing either λ_r or λ_i strengthens the departures from Gaussianity, suggesting that the lower frequency and least-damped barotropic eigenmodes could have the greatest non-Gaussian behavior. Changing these parameters can also change the covariance between the real and imaginary parts of $\boldsymbol{\Psi}$, and thus the rotation applied in Fig. 8.

TABLE 2. Skewness and excess kurtosis of the marginal distributions $p(\Psi_r)$ and $p(\Psi_i)$ of the stochastic Rossby wave model with a stochastic basic state. The excess kurtosis measures the departure from Gaussianity: excess kurtosis = kurtosis −3. Note that the distributions are nearly Gaussian, even for strong multiplicative noise.

	Skewness		Excess kurtosis	
Multiplicative noise σ_r^M	Ψ_r	Ψ_i	Ψ_r	Ψ_I
0.0	0.0	0.0	0.0	0.0
0.2	0.0	0.0	0.0	0.0
0.4	0.0	−0.1	0.1	0.0
0.6	−0.1	−0.2	0.3	0.2
0.8	−0.2	−0.4	0.9	0.9

JOBNAME: XJAS 62#2 2005 PAGE: 11 SESS: 69 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

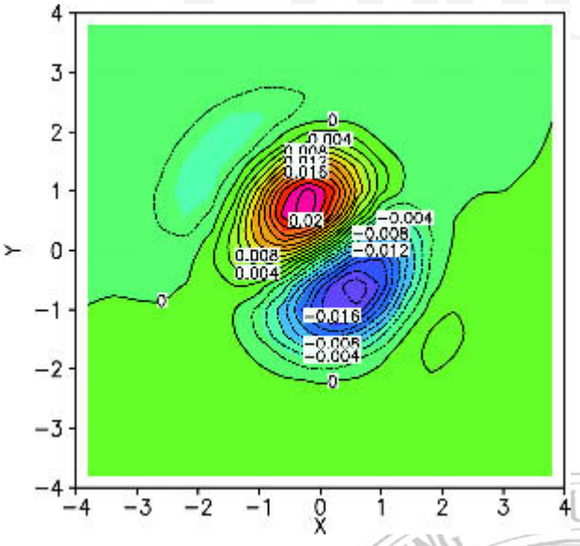

FIG. 8. Steady state PDF anomalies (departures from bivariate Gaussianity) of the Rossby wave model ($\lambda_r = 0.07$ and $\lambda_i = 0.19$) with a stochastic frictional damping ($\sigma_r^M = 0.05$). The axes are rotated so that $X \equiv 0.86\Psi_r - 0.51\Psi_i$ and $Y \equiv 0.5\Psi_r - 0.86\Psi_i$. The contour interval is 0.002.

5. Observed atmospheric PDFs

Having demonstrated that non-Gaussian behavior can result from simple linear multiplicative noise, we next compare the observed non-Gaussianity to the multiplicative stochastic paradigm. Following previous studies (e.g., Mo and Ghil 1988; Molteni et al. 1990; Kimoto and Ghil 1993a,b; Corti et al. 1999; Smyth et al. 1999; Weisheimer et al. 2001), we consider non-Gaussian regimes in a highly truncated bivariate phase space spanned by two leading EOFs.

5.a. Bivariate PDF

The analysis is applied to Northern Hemisphere 750-hPa streamfunction data for the extended winters (November–March) of 1949/50–2001/02, spectrally truncated to T21 resolution. The data were obtained from the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP/NCAR) Reanalysis dataset (Kalnay et al. 1996). Streamfunction anomalies were defined by removing the seasonal cycle (i.e., the annual mean plus the first three annual harmonics) from each variable at each gridpoint and then applying a 7-day running mean filter. Calculations using 500-hPa geopotential height pentad data yielded similar results (not shown).

A principal component analysis (PCA) was applied to the streamfunction anomalies. The first two EOF patterns form the orthogonal basis vectors of a reduced phase space. The first EOF (EOF1 with principal component PC1) explains 16.6% of the total variance, whereas the second EOF (EOF2 with principal component PC2) explains 8.2%. The EOFs (not shown) resemble those found in earlier studies (e.g., Kimoto and Ghil 1993a; Smyth et al. 1999); signs are defined as in Kimoto and Ghil (1993a). The PCs are normalized to have zero mean and unit standard deviation. Monahan et al. (2003) suggested that the leading atmospheric regimes have a large projection on PC1 and PC2 and a small, but non-negligible, projection on PC3. Including this projection on PC3 did not affect any of the results shown below.

The bivariate PDF of PC1 and PC2, determined by dividing the interval [−4: 4, −4: 4] into 20 × 20 equal bins and then applying a 3 × 3 bin smoothing, is shown in Fig. 9a. Although the PDF is generated with a fairly simple technique, virtually identical results are obtained using bivariate Gaussian kernel density estimation with optimal bandwidth $h = 0.25$ (not shown). Positive (negative) departures from Gaussianity (Fig. 9b) indicate that the observed PDF is greater (smaller) than the corresponding bivariate Gaussian distribution.

Past studies of bivariate PDFs such as those shown here tend to focus on the local maxima, determined for example using some bump-hunting algorithm, identifying them as regimes and producing corresponding composite anomaly maps. Such maxima are, however, sensitive to the dataset used, to the smoothing applied to the PDF, and perhaps most of all to sampling considerations (particularly the data period). For example, the maximum at (0, −2.5) lies in different locations when only the years 1949–75 or only the years 1976–2001 are used, a maximum at (1, 1.5) is relatively more prominent for the years 1958–98 (e.g., Monahan et al. 2001).

Considerably more robust across all studies, including this one, is the overall pattern of the departures from bivariate Gaussianity, consisting of two slightly arcing ridge/trough pairs aligned roughly along a line from the upper left quadrant to the lower right quadrant. Notably, the bivariate skewness (e.g., Stephenson et al. 2004) is 0.28, which is 99% significant, whereas excess bivariate kurtosis (0.36) is 90% significant. Note also the pronounced negative region, this feature has an amplitude as large as the positive departures, yet has generally received far less attention. The alternating positive/negative departures pattern exists in many different subsamples of the data using only the odd years, even years, or the years 1949–75, or separately the years 1976–2001. Also, broad regions of positive and negative departures are statistically significant at the 90% confidence level (heavy contours in Fig. 9a) determined using the Monte Carlo method employed in Kimoto and Ghil (1993a).

5.b. Balancing the probability budget

Using the Fokker–Planck Eq. (4), a steady climatological probability budget for a bivariate system ($i = 1, 2$; $j = 1, 2$), assumed to be Markovian, is

JOBNAME: XJAS 62#2 2005 PAGE: 12 SESS: 68 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

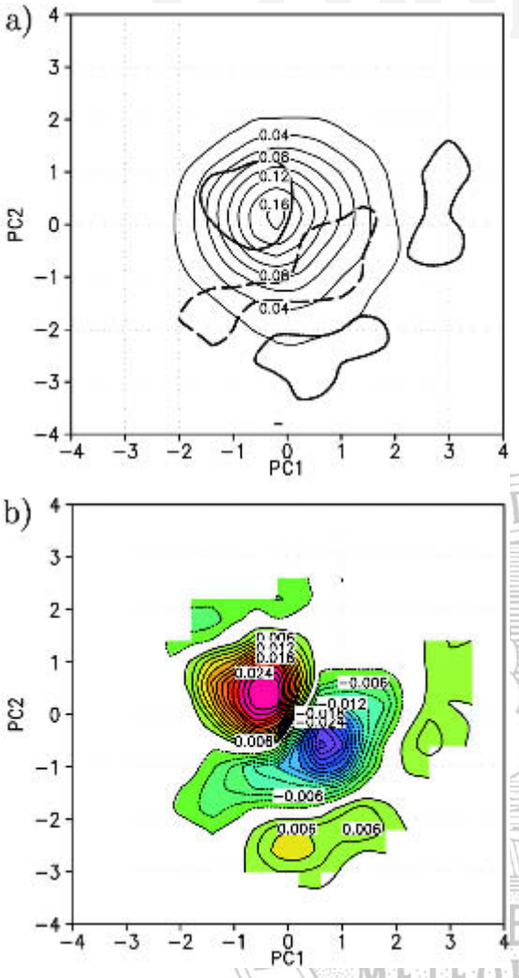

FIG. 9. (a) PDF of PC1 and PC2 (thin solid lines). Regions of significant positive deviations from a bivariate Gaussian PDF are indicated by thick solid lines, significant regions of negative deviations are indicated by thick dashed lines. The deviations from a bivariate Gaussian PDF are significant at the 90% confidence level. (b) The actual PDF anomalies of PC1 and PC2. The zero contour is omitted for clarity, the contour interval is 0.002.

$$-\sum_i \frac{\partial}{\partial x_i} A_i p(\mathbf{x}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\mathbf{B}\mathbf{B}^{\mathrm{T}})_{ij}\, p(\mathbf{x}, t) = 0. \tag{18}$$

That is, in a time-averaged sense, the effective drift balances the stochastic diffusion. We can then use Eq. (18) to ask whether the nonlinear effective drift due to interaction between PC1 and PC2, estimated by using its finite difference approximation (6), is sufficient to produce the non-Gaussianity of the observed PDF. That is, we ask whether or not the probability budget can be balanced using purely additive stochastic forcing.

The effective drift (and the related standard errors) of the data components are shown in Fig. 10. The uncertainties are relatively small, the standard error is about 0.01–0.02 for most of the points. Note that the drift is nearly linear. F10

At this point it is important to test if the white-noise approximation made in the bivariate version of Eq. (2) is justifiable for the data used. For any observation $\mathbf{x}_{\mathrm{obs}}(t)$ the effective drift $\mathbf{A}(\mathbf{x})$ estimated from data (see Fig. 10) can be used to calculate a forward time step Δt: $\mathbf{x}(t + \Delta t) = \mathbf{A}(\mathbf{x})\Delta t + \mathbf{x}_{\mathrm{obs}}(t)$. If the white-noise assumption is correct the residual

$$\mathbf{r} \equiv \mathbf{x}_{\mathrm{obs}}(t + \Delta t) - \mathbf{x}(t + \Delta t) \tag{19}$$

should equal the multiplicative white-noise term $\mathbf{B}(\mathbf{x})\eta$. Indeed, the residual is almost uncorrelated on the resolved time scale [i.e., $\mathbf{r}(t = 7\ \mathrm{days}) \approx 0$, as shown in Fig. 11]. In contrast, the decorrelation time scales of PC1 and PC2 are about 12 and 15 days and, therefore, the white-noise approximation can be justified. F11

To the extent that the white-noise approximation is valid, we ask whether or not the probability budget can be balanced using pure additive noise. That is, if we rewrite Eq. (18) for the special case where $\mathbf{B}$ is purely additive ($i = 1, 2$; $j = 1, 2$),

$$-\sum_i \frac{\partial}{\partial x_i} A_i p(\mathbf{x}, t) + \frac{1}{2} \sum_{i,j} (\mathbf{B}_A \mathbf{B}_A^{\mathrm{T}})_{ij} \frac{\partial^2}{\partial x_i \partial x_j} p(\mathbf{x}, t) = R, \tag{20}$$

is the residual R negligibly small? Because the errors in a direct estimation of the diffusion term are large (Sura and Barsugli 2002), we instead use the effective drift $\mathbf{A}(\mathbf{x})$ (Fig. 10) and the steady-state FDR (5) to first determine the covariance of the noise $\langle \mathbf{B}(\mathbf{x})\mathbf{B}^{\mathrm{T}}(\mathbf{x}) \rangle$ required to balance the drift in (18):

$$\langle \mathbf{B}(\mathbf{x})\mathbf{B}^{\mathrm{T}}(\mathbf{x}) \rangle = \begin{pmatrix} 0.079 & 0.001 \\ 0.001 & 0.103 \end{pmatrix}. \tag{21}$$

Next, using the PDF (Fig. 9), the effective drift (Fig. 10), and assuming that the noise given by the matrix (21) is purely additive (i.e., assuming $\mathbf{B}_A\mathbf{B}_A^{\mathrm{T}} = \langle \mathbf{B}(\mathbf{x})\mathbf{B}^{\mathrm{T}}(\mathbf{x}) \rangle$) the two terms on the left-hand side of the Fokker–Planck Eq. (20) are evaluated and shown in Figs. 12a and 12b. Their sum (the residual R) is shown in Fig. 12c. The residual is not zero, indicating that pure additive noise is not able to balance the probability budget. F12

Despite the relative shortness of the data record, this result appears robust. For example, it is virtually unchanged if the data subsamples listed at the end of section 5.a are used. Moreover, the residual is statistically significant at the 90% confidence level, determined as before, using a Monte Carlo method. One thousand 50-yr time series having the same zero-lag covariances as the original time series were generated by using the effective drift (Fig. 10) and the corresponding additive noise [Eq. (21)], and the deterministic and additive sto-

JOBNAME: XJAS 62#2 2005 PAGE: 13 SESS: 68 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

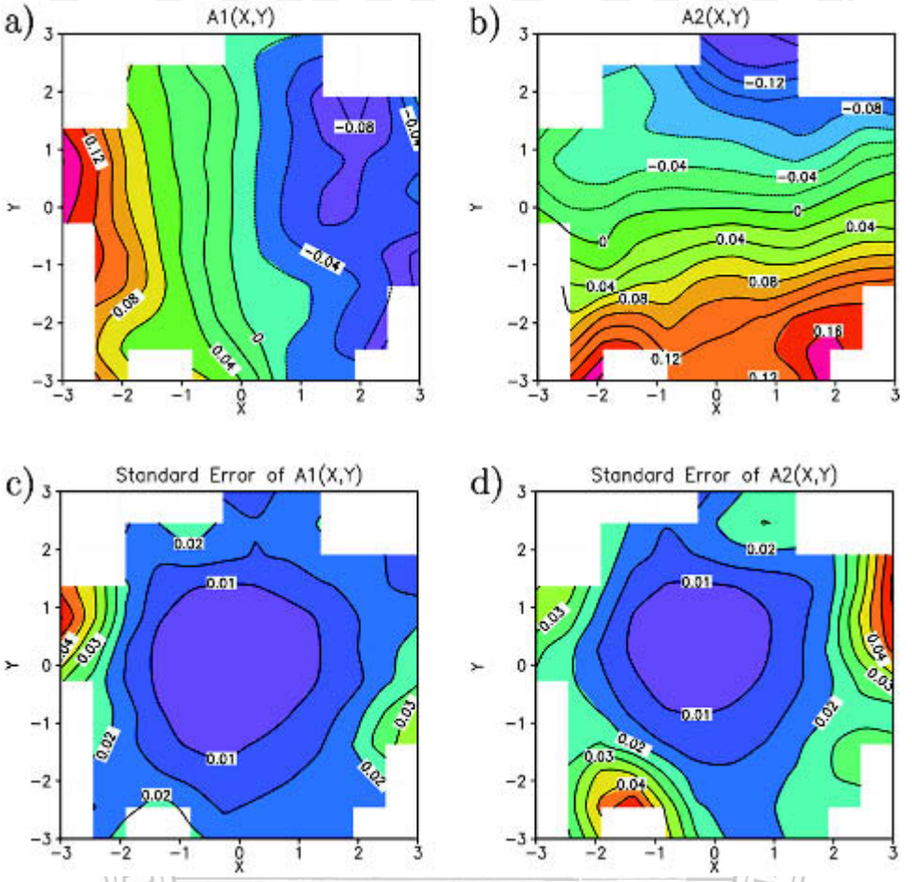

FIG. 10. The effective drift (and its standard error) estimated from data. (a) $A1(x,y)$ denotes the x component (PC1), and (b) $A2(x,y)$ denotes the y component (PC2) of the two-dimensional system. The contour and shading interval is 0.02. (c),(d) Related standard errors with a contour and shading interval 0.005.

chastic terms (and their sum) in the Fokker–Planck equation were computed for each of them. The number of those random terms larger/smaller than the original terms were used to obtain confidence intervals at each point in the bivariate phase space.

If we partition the net noise effect $\mathbf{B}\mathbf{B}^{\mathrm{T}}$ in (18) into an additive noise term $\mathbf{B}_A\mathbf{B}_A^{\mathrm{T}}$ and a multiplicative noise term $\mathbf{B}_R\mathbf{B}_R^{\mathrm{T}}$ then $-R$ would represent the multiplicative part:

$$R = -\frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}(\mathbf{B}_R\mathbf{B}_R^{\mathrm{T}})_{ij}\,p(\mathbf{x}, t), \qquad (22)$$

with R and $\mathbf{B}_A\mathbf{B}_A^{\mathrm{T}}$ appropriately scaled so that the fluctuation–dissipation relation is satisfied.

The key point here is that if the departures from Gaussianity were primarily due to the nonlinear drift term, then the probability budget for the joint PC1/2 PDF would be balanced with pure additive noise. Clearly, it is not; all three terms of the Fokker–Planck

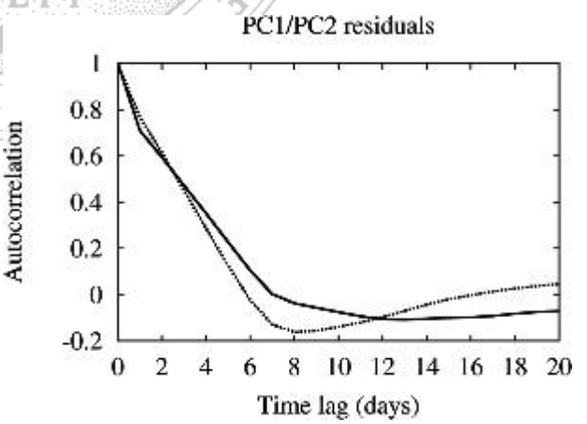

FIG. 11. Autocorrelation functions of the residual $\mathbf{r} \equiv \mathbf{x}_{\mathrm{obs}}(t + \Delta t) - \mathbf{x}(t + \Delta t)$ for PC1 (solid line) and PC2 (dashed line). Note that for both PC1 and PC2 the autocorrelation is close to zero after the resolved time scale of seven days; see section 5.b for details.

JOBNAME: XJAS 62#2 2005 PAGE: 14 SESS: 68 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

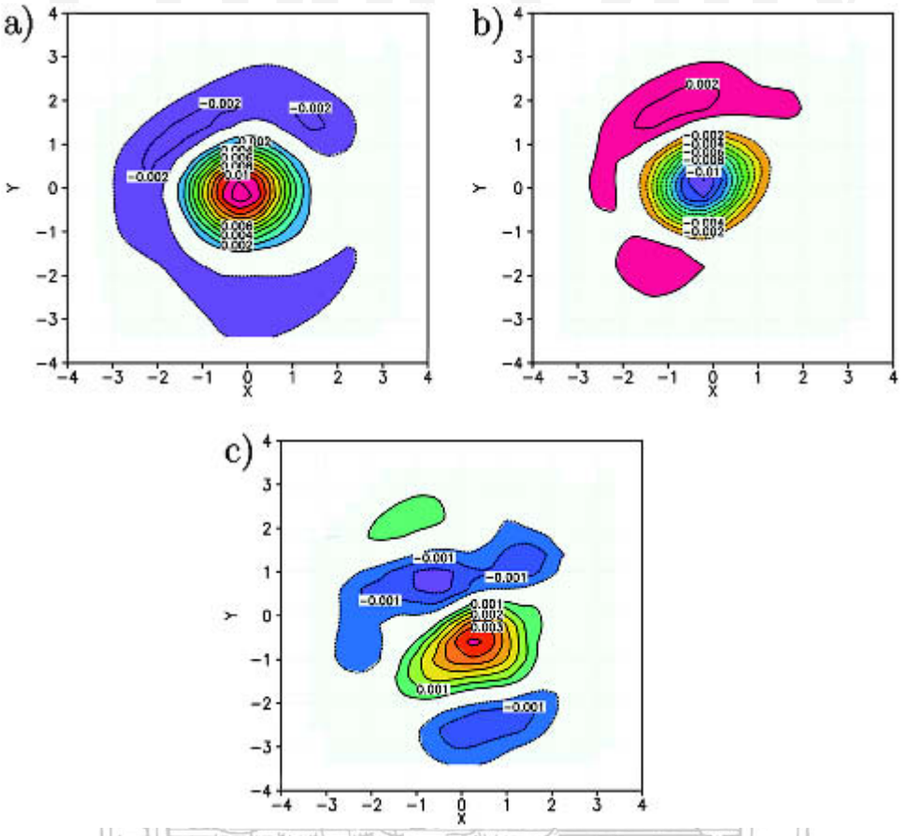

FIG. 12. (a) The effective deterministic term (C.I. = 0.001), (b) the stochastic term (C.I. = 0.001, the dark dashed line indicates the zero contour; see text for details), and (c) the sum of the effective deterministic and the stochastic term of the steady Fokker–Planck Eq. (20) evaluated by using the PDF (Fig. 9), the deterministic drift (Fig. 10), and pure additive noise given by the matrix (21): C.I. = 0.0005. The negative value of that sum ($-R$) can be interpreted as the multiplicative noise contribution of the Fokker–Planck Eq. (20), positive (negative) values mean that the additive noise is too weak (strong) to balance the deterministic term. The shown regions are significant at the 90% confidence level. The x component represents PC1, the y component represents PC2.

equation have the same order of magnitude. Positive (negative) values in Fig. 12c mean that the additive noise is too weak (strong) to balance the drift. Thus, the observed departures from Gaussianity do not result from the nonlinear drift term defined in the two-dimensional space, but rather from the multiplicative structure of the noise.

6. Summary and discussion

In this paper we outlined a stochastic perspective on atmospheric regime behavior based on a treatment of climate variability as a stochastic system with state-dependent noise. We demonstrated how some simple linear (or nearly linear) systems with multiplicative noise can produce non-Gaussian regime-like behavior without multiple equilibrium solutions of the governing equations. The presence of non-Gaussianity, therefore, does not by itself imply that a system has deterministic nonlinear multiple regimes, nor that these regimes have a noticeable enhancement of persistence or predictability due to this nonlinearity.

The linear Rossby wave response to steady forcing is non-Gaussian when the wave damping has a stochastic component, as might be expected, for example, from gustiness in boundary layer dissipation. This multiplicative noise leads to a non-Gaussian distribution due to an intermittent behavior of the Rossby waves. Many nonlinear systems are intermittent in time, space, or both (e.g., Sreenivasan 1999; Sreenivasan and Antonia 1997). Again, the most common explanation for intermittency is that it is induced by nonlinearities in the slow manifold of the governing equations.

One motivation for this paper is our earlier work with linear inverse modeling (LIM; Winkler et al. 2001; Newman et al. 2003), in which observed weekly vari-

JOBNAME: XJAS 62#2 2005 PAGE: 15 SESS: 70 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

ability of 250- and 750-hPa streamfunction is successfully modeled and predicted with a multivariate linear system [with $O(30)$ degrees of freedom] plus noise. Interestingly, both predictability and forecast skill (for example, wintertime week-3 skill, shown with shading in Fig. 13) is high in regions where non-Gaussianity is relatively large (cross-hatching in Fig. 13). In fact, the deterministic linear operator in LIM cannot distinguish between a linear system driven by additive noise and one that also includes linear multiplicative noise, because then the noise-induced drift is part of the linear operator returned by LIM. In the former case, LIM returns the deterministic linear operator $\mathbf{L}$, and the "best" (in the least squares error sense) forecasts of $\mathbf{x}$ at forecast lead τ are $\mathbf{x}(t + \tau) = \mathbf{G}(\tau)\mathbf{x}(t) = \exp(\mathbf{L}\tau)\mathbf{x}(t)$. In the latter case, LIM returns the effective linear operator $\tilde{\mathbf{L}} = \mathbf{L} + (1/2)\mathbf{B}^2$ (i.e., the noise-induced drift is part of the linear operator returned by LIM), and the best forecasts are $\mathbf{x}(t + \tau) = \mathbf{G}(\tau)\mathbf{x}(t) = \exp(\tilde{\mathbf{L}}\tau)\mathbf{x}(t)$. Trying to differentiate between these two models is a major focus of our current research.

In light of the success of LIM, the similarity between Figs. 8 and 9b and the analysis of the probability budget of the bivariate PC1/2 PDF suggests that atmospheric variability may be represented by a phase space that can be characterized not so much by localized regimes as by less-localized departures from Gaussianity that are largely a result of an effectively almost linear system perturbed by multiplicative noise. However, while our observational results are suggestive, they do not constitute a proof. Notably, we do not yet have a solution for what $\mathbf{B}(\mathbf{x})$ should be for the real atmosphere, either by determining it from the extremely limited dataset, from extended model datasets, or from a simplification of the equations of motion.

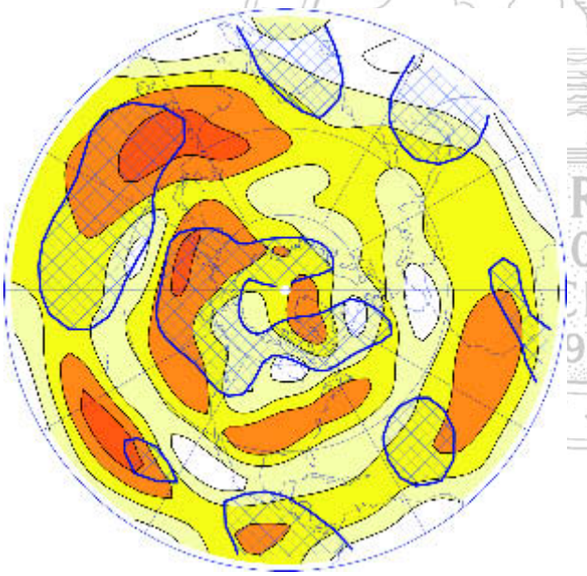

FIG. 13. Skill of the LIM compared with regions of non-Gaussianity, adopted from Newman et al. (2003). Shading represents local anomaly correlation of week-3 250-hPa streamfunction observations with LIM forecasts for Dec–Feb 1969/70–2000/01. Contour interval is 0.1. Shading of positive values starts at 0.2, darker shading denotes larger correlation values. Cross-hatching represents region of significant non-Gaussianity, defined as locations where the value of the nonparametric two-sample Kolmogorov–Smirnov measure indicates that the observed PDF is significantly different from Gaussianity at the 95% confidence level [Kendall and Stuart (1977); also see the appendix in Sardeshmukh et al. (2000)].

All complete models of regime behavior must deal with the closure problem. That is, there must not only be a deterministic model [e.g., $\mathbf{N_1}$ in Eq. (1)] of the resolved anomalies, but also consideration of the effects of the unresolved anomalies [e.g., $\mathbf{N_2}$ and $\mathbf{N_3}$ in Eq. (1)]. For low-frequency variability, this partition may not be well defined, especially since there is not a clean separation of time scales in the atmosphere. Thus, the two paradigms of non-Gaussian atmospheric behavior discussed here need not be mutually exclusive. For example, some or even all of what appears as multiplicative noise in a severely truncated system such as the PC1/PC2 space could appear as deterministic nonlinear interactions in a higher-order space. Given that $\mathbf{r}$ [Eq. (19)] has a relatively short decorrelation time, however, it is reasonable to suggest that, on the weekly time scale, the multiplicative noise closure results in as much understanding of the non-Gaussianity of PC1 and PC2 as would result from including many more degrees of freedom in the resolved portion of the flow. And in a practical sense, of course, all of these degrees of freedom must be simulated and predicted, which might still require a multiplicative noise closure for their understanding.

Some other studies (e.g., Achatz and Branstator 1999; D'Andrea and Vautard 2001) have modeled nonlinear baroclinic atmospheric model output using a relatively low-order [$O(30)$] empirical nonlinear deterministic model, although it is not clear how important the nonlinearity is for reproducing the statistics of variability as opposed to reproducing the correct mean climate. Nevertheless, our results suggest that a comprehensive approach toward a better understanding of atmospheric regime behavior must consider state-dependent noise. Further research is required to assess the more detailed extent to which multiple non-Gaussian regimes may be due to the nature of unpredictable stochastic forcing, rather than to the slow, predictable deterministic nonlinear dynamics of the atmosphere.

Acknowledgments. The authors wish to thank Adam Monahan and two anonymous reviewers for comments that improved this paper. This work was funded by the Predictability DRI of the Office of Naval Research, Grant N00014–99–1-0021.

JOBNAME: XJAS 62#2 2005 PAGE: 16 SESS: 71 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

APPENDIX A

Fitting a PDF to deterministic and stochastic models

In this appendix we show how various combinations of different deterministic and stochastic terms in the governing SDE can give rise to a given PDF. We consider a univariate Itô SDE of the form

$$dx = A(x)dt + B(x)dW, \tag{A.1}$$

where $A(x)$ and $B(x)$ are sufficiently smooth and bounded functions, and W denotes a Wiener process. The PDF $p(x, t)$ of the Itô SDE (A.1) is governed by the corresponding Itô Fokker–Planck equation (e.g., Gardiner 1985; Horsthemke and Léfèver 1984; Paul and Baschnagel 1999):

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}A(x)p(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}B(x)^2 p(x,t). \tag{A.2}$$

Given a stationary non-Gaussian PDF $p(x)$ (with $p(-\infty) = p(\infty) = 0$) and pure additive noise $B^2 \equiv$ const., the stationary Fokker–Planck equation can be solved for $A(x)$:

$$A(x) = \frac{B^2}{2}\frac{\partial \ln p(x)}{\partial x} + \frac{\beta}{p(x)}, \tag{A.3}$$

with the arbitrary constant β. Given the same PDF $p(x)$ and the deterministic damping term $A(x)$, the stationary Fokker–Planck equation can be solved for $B(x)$:

$$B(x) = \left(\frac{2}{p(x)}\int_{-\infty}^{x}[A(x')p(x') + \beta]dx'\right)^{1/2}, \tag{A.4}$$

with an arbitrary constant β. That means that the same PDF can either be produced by nonlinear deterministic dynamics with additive noise, or, if $A(x) = -\varepsilon x$, by linear deterministic dynamics with multiplicative noise. Because we expect $A(x)$ and $B(x)$ to be bounded in physically reasonable situations, the constant is set to zero: $\beta = 0$.

Even if deterministic and stochastic regimes do have the same stationary PDFs, the dynamical properties (and, therefore, the predictabilities) of the trajectories are fundamentally different. This behavior can be illustrated by the mean residence times. The residence time of a stochastic trajectory initially at x inside the interval $[x_1, x_2]$ is the time until the trajectory first hits the boundary of the interval. The mean residence time $\tau(x)$ of a stochastic trajectory governed by the SDE (A.1) is given by (e.g., Gardiner 1985; Horsthemke and Léfèver 1984; Paul and Baschnagel 1999),

$$A(x)\frac{\partial \tau}{\partial x} + \frac{1}{2}B(x)^2\frac{\partial^2 \tau}{\partial x^2} + 1 = 0, \tag{A.5}$$

with the boundary conditions $\tau(x_1) = \tau(x_2) = 0$.

APPENDIX B

An analytical solution of a Fokker–Planck equation with multiplicative noise

In this appendix an analytical solution of the Itô and Stratonovich Fokker–Planck equation for a stochastic basic state and no steady Rossby wave forcing [$\mathbf{S} = (0,0)^{\mathrm{T}}$] is presented. This solution is then used to discuss the fundamental physical difference between a stochastic frictional damping and a stochastic basic state. The SDE for the time evolution of the vector $\boldsymbol{\Psi} = (\Psi_r, \Psi_i)^{\mathrm{T}}$ is

$$\frac{d\boldsymbol{\Psi}}{dt} = \mathbf{A}\boldsymbol{\Psi} + \mathbf{B}\boldsymbol{\eta}^M + \boldsymbol{\eta}^A, \tag{B.1}$$

with the matrices

$$\mathbf{A} = \begin{pmatrix} -r_0 & -mD_n \\ mD_n & -r_0 \end{pmatrix} \tag{B.2}$$

and

$$\mathbf{B} = \begin{pmatrix} -\Psi_i & 0 \\ \Psi_r & 0 \end{pmatrix}, \tag{B.3}$$

and the multiplicative and additive stochastic forcing vectors $\boldsymbol{\eta}^M = (\eta_r^M, \eta_i^M)^{\mathrm{T}}$, and $\boldsymbol{\eta}^A = (\eta_r^A, \eta_i^A)^{\mathrm{T}}$. The stochastic components $\eta_r^M, \eta_i^M, \eta_r^A$, and η_i^A are assumed to be independent Gaussian white-noise processes with corresponding amplitudes $\sigma_r^M, \sigma_i^M, \sigma_r^A$, and σ_i^A [Eq. (16)]. Note that because of the structure of the matrix (B.3) the imaginary part of the multiplicative noise η_i^M (with amplitude σ_i^M) has no impact and is, therefore, not specified. For identical additive noise amplitudes $\sigma_r^A = \sigma_i^A = \sigma^A$ the corresponding Fokker–Planck equation for the PDF $p(\boldsymbol{\Psi}, t)$ reads ($i = 1,2$; $j = 1,2$):

$$\begin{aligned}\frac{\partial p(\boldsymbol{\Psi},t)}{\partial t} = &-\sum_i \frac{\partial}{\partial \Psi_i}\Bigg[(\mathbf{A}\boldsymbol{\Psi})_i \\ &+ \alpha\sum_{j,k}(\sigma_r^M)^2\left(\frac{\partial}{\partial \Psi_j}B_{ik}\right)B_{jk}\Bigg]p(\boldsymbol{\Psi},t) \\ &+ \frac{1}{2}\sum_{i,j}(\sigma_r^M)^2\frac{\partial^2}{\partial \Psi_i \partial \Psi_j}(\mathbf{B}\mathbf{B}^{\mathrm{T}})_{ij}\,p(\boldsymbol{\Psi},t) \\ &+ \frac{1}{2}\sum_i(\sigma^A)^2\frac{\partial^2}{\partial \Psi_i^2}p(\boldsymbol{\Psi},t),\end{aligned} \tag{B.4}$$

with $\alpha = 0$ for Itô systems and $\alpha = 1/2$ for Stratonovich systems. Because of the identical additive noise amplitudes $\sigma_r^A = \sigma_i^A = \sigma^A$ the Fokker–Planck equation has an exact stationary Gaussian solution even in the presence of multiplicative noise with amplitude σ_r^M:

JOBNAME: XJAS 62#2 2005 PAGE: 17 SESS: 70 OUTPUT: Mon Jan 10 12:49:54 2005
/Pages/ams/xjas/89445/jas3408

$$p(\boldsymbol{\Psi}) = \Phi \exp[-\gamma(\Psi_r^2 + \Psi_i^2)], \tag{B.5}$$

with the normalization constant Φ, and

$$\gamma = \left(\frac{r_0 - \frac{1}{2}(\sigma_r^M)^2 + \alpha(\sigma_r^M)^2}{(\sigma^A)^2} \right). \tag{B.6}$$

Note that the noise-induced drift in the Stratonovich system ($\alpha = 1/2$) effectively increases the damping of the mean flow. Yet, the noise-induced drift is compensated by the remaining effect of the multiplicative noise. Thus, in the Stratonovich calculus there is no effect of the multiplicative noise on the stationary PDF at all.

This behavior can be understood by simple geometric considerations described below. These considerations also clarify the fundamental physical difference between a stochastic frictional damping and a stochastic basic state. A schematic explanation is shown in Fig. B.1. The equation of a circle (the undamped motion of the Rossby wave with or without the presence of additive noise with identical amplitudes $\sigma_r^A = \sigma_i^A = \sigma^A$) with radius a around the origin of the coordinate system in the (Ψ_r, Ψ_i) plane is

$$f(\Psi_r, \Psi_i) = \Psi_r^2 + \Psi_i^2 = a^2. \tag{B.7}$$

A vector perpendicular to the circle (B.7) is $\nabla f(\Psi_r, \Psi_i) \propto (-\Psi_r, -\Psi_i)$. A vector tangential to the circle (B.7) is $\nabla f(\Psi_r, \Psi_i) \times e_z \propto (-\Psi_i, \Psi_r)$, where e_z is the unit vector perpendicular to the (Ψ_r, Ψ_i) plane. It is clear from Eqs. (14) and (15), and the multiplicative noise vector $\boldsymbol{\eta}^A = (\eta_r^A, \eta_i^A)^T$, that the kicks of the stochastic damping are perpendicular to the undamped circular phase space motion, whereas the stochastic kicks in the basic state are tangential to it. Therefore, the stochasticity in the zonal basic state u_0 only influences the phase of the Rossby wave, but not its energy.

When stochastic perturbations occur in the frictional damping r_0, the energy is altered, but not the phase of the Rossby wave. In Fig. B.1 the thick solid circle indicates the energy conserving wave motion of the undamped system with or without the presence of additive noise with identical amplitudes $\sigma_r^A = \sigma_i^A = \sigma^A$. The thick dashed line shows the phase space motion of the corresponding damped system. The thin dashed circles are isolines of the strength [defined as $(\sigma_r^M \Psi_r)^2 + (\sigma_r^M \Psi_i)^2$] of the multiplicative stochastic forcing. The arrows show the directions of the stochastic kicks in (a) the zonal basic state and (b) the frictional damping.

Because the stochasticity in the basic state only influences the phase of the Rossby wave, the PDF of the system remains strictly Gaussian for that kind of multiplicative noise. Yet, since the multiplicative noise in the frictional damping alters the energy of the wave, the PDF becomes highly non-Gaussian.

The situation changes if the undamped trajectory does not coincide with the circular lines of constant strength of the multiplicative stochastic forcing. This situation occurs if nonidentical additive noise amplitudes ($\sigma_r^A \neq \sigma_i^A$) are used, or if a steady Rossby wave forcing [$\mathbf{S} \neq (0, 0)^T$] is imposed on the governing Eq. (B.1). Then, the stationary PDF becomes slightly non-Gaussian even in the case of a stochastic basic state. Nonetheless, the effect of a stochastic frictional damping on the non-Gaussianity of the systems PDF is much stronger than the corresponding effect of a stochastic basic state (see Fig. 6).

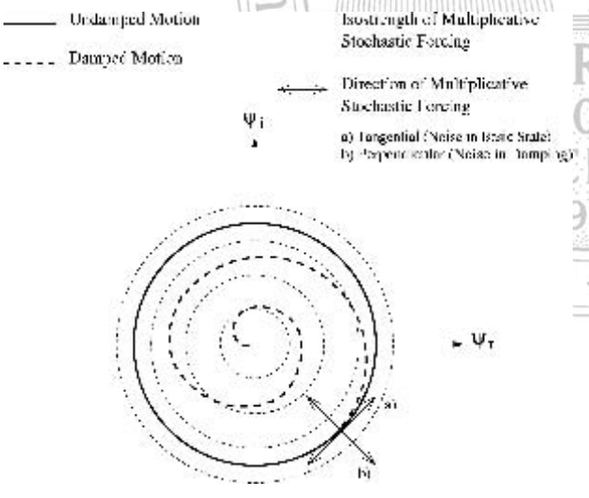

FIG. B1. A schematic sketch to show the fundamental physical difference between a stochastic frictional damping and a stochastic basic state. The thick solid circle shows the energy conserving wave motion of the undamped system. The thick dashed line shows the phase space motion of the corresponding damped system. The thin dashed circles are isolines of the strength [defined as $(\sigma_r^M \Psi_r)^2 + (\sigma_r^M \Psi_i)^2$] of the multiplicative stochastic forcing. The arrows show the directions of the stochastic kicks in (a) the zonal basic state (tangential) and (b) the frictional damping (perpendicular). See appendix B for a more detailed discussion.

REFERENCES

Achatz, U., and G. Branstator, 1999: A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability. *J. Atmos. Sci.,* **56,** 3140–3160.

Baur, F., 1947: *Musterbeispiele Europäischer Großwetterlagen.* Dieterich, xx pp.

——, F. Hess, and P. Nagel, 1944: *Kalender der Großwetterlagen Europas 1881–1939.* Forschungsinstitut für langfristige Witterungsvorhersage, xx pp.

Borges, M. D., and P. D. Sardeshmukh, 1995: Barotropic Rossby-wave dynamics of zonally varying upper-level flows during northern winter. *J. Atmos. Sci.,* **52,** 3779–3796.

Buizza, R., M. J. Miller, and T. N. Palmer, 1999: Stochastic simulation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **125,** 2887–2908.

Cash, B. A., and S. Lee, 2001: Observed nonmodal growth of the Pacific–North American teleconnection pattern. *J. Climate,* **14,** 1017–1028.

Chang, J. S., and G. Cooper, 1970: A practical difference scheme for Fokker–Planck equations. *J. Comput. Phys.,* **6,** 1–16.

Charney, J. G., and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.,* **36,** 1205–1216.

Cheng, X., and J. M. Wallace, 1993: Cluster analysis of Northern Hemisphere wintertime 500-hPa height field. *J. Atmos. Sci.,* **50,** 2674–2696.

Corti, S., F. Molteni, and T. N. Palmer, 1999: Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature,* **29,** 799–802.

D'Andrea, F., and R. Vautard, 2001: Extratropical low-frequency variability as a low-dimensional problem. I: A simplified model. *Quart. J. Roy. Meteor. Soc.,* **127,** 1357–1374.

DelSole, T., and B. F. Farrell, 1995: A stochastically excited linear system as a model for quasigeostrophic turbulence: Analytic results for one- and two-layer fluids. *J. Atmos. Sci.,* **52,** 2531–2547.

DeSwart, H. E., 1988: Low-order spectral models of the atmospheric circulation: A survey. *Acta Appl. Math.,* **11,** 49–96.

Egger, J., 1981: Stochastically driven large-scale circulations with multiple equilibria. *J. Atmos. Sci.,* **38,** 2606–2618.

Ewald, B., C. Penland, and R. Temam, 2004: Accurate integrations of stochastic climate models. *Mon. Wea. Rev.,* **132,** 154–164.

Farrell, B. F., and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.,* **53,** 2025–2040.

Friedrich, R., and Coauthors, 2000: Extracting model equations from experimental data. *Phys. Lett. A,* **271,** 217–222.

Gardiner, C. W., 1985: *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Science.* 2nd ed. Springer Verlag, 442 pp.

Ghil, M., and S. Childress, 1987: *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics.* Springer Verlag, 485 pp.

——, and A. W. Robertson, 2002: "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? *Proc. Natl. Acad. Sci. U.S.A,* **99,** 2493–2500.

Gradišek, J., S. Siegert, R. Friedrich, and I. Grabec, 2000: Analysis of time series from stochastic processes. *Phys. Rev. E,* **62,** 3146–3155.

Hansen, A. R., and A. Sutera, 1986: On the probability density distribution of planetary scale atmospheric wave amplitude. *J. Atmos. Sci.,* **43,** 3250–3265.

Hart, J., 1979: Barotropic quasigeostrophic flow over anisotropic mountains. *J. Atmos. Sci.,* **36,** 1736–1746.

Hasselmann, K., 1976: Stochastic climate models. Part I. Theory. *Tellus,* **28,** 473–484.

Horsthemke, W., and R. Léfèver, 1984: *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology.* Springer Verlag, 318 pp.

Hoskins, B. J., A. J. Simmons, and D. G. Andrews, 1977: Energy dispersion in a barotropic atmosphere. *Quart. J. Roy. Meteor. Soc.,* **103,** 553–567.

Itoh, H., and M. Kimoto, 1996: Multiple attractors and chaotic itinerancy in a quasigeostrophic model with realistic topography: Implications for weather regimes and low-frequency variability. *J. Atmos. Sci.,* **53,** 2217–2231.

——, and ——, 1997: Chaotic itinerancy with preferred transition routes appearing in an atmospheric model. *Physica D,* **109,** 274–292.

——, and ——, 1999: Weather regimes, low-frequency oscillations, and principal patterns of variability: A perspective of extratropical low-frequency variability. *J. Atmos. Sci.,* **56,** 2684–2705.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kendall, M., and A. Stuart, 1977: *The Advanced Theory of Statistics.* Vol. 1, 3rd ed. Macmillan, 283 pp.

Khasminskii, R. Z., 1966: A limit theorem for the solutions of differential-equations with random right hand side. *Theory Probab. Its Appl.,* **11,** 390–406.

Kimoto, M., and M. Ghil, 1993a: Multiple flow regimes in the Northern Hemisphere winter. Part I: Methodology and hemispheric regimes. *J. Atmos. Sci.,* **50,** 2625–2643.

——, and ——, 1993b: Multiple flow regimes in the Northern Hemisphere winter. Part II: Sectorial regimes and preferred transitions. *J. Atmos. Sci.,* **50,** 2645–2672.

Kloeden, P., and E. Platen, 1992: *Numerical Solution of Stochastic Differential Equations.* Springer Verlag, 632 pp.

Landa, P. S., and P. V. E. McClintock, 2000: Changes in the dynamical behavior of nonlinear systems induced by noise. *Phys. Rep.,* **323,** 1–80.

Legras, B., and M. Ghil, 1985: Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci.,* **42,** 433–471.

Lin, J. W. B., and J. D. Neelin, 2002: Considerations for stochastic convective parameterization. *J. Atmos. Sci.,* **59,** 959–975.

Majda, A. J., I. Timofeyev, and E. Vanden Eijnden, 1999: Models for stochastic climate prediction. *Proc. Natl. Acad. Sci. U.S.A.,* **96,** 14 687–14 691.

——, ——, and ——, 2003: Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.,* **60,** 1705–1722.

Mo, K., and M. Ghil, 1988: Cluster analysis of multiple planetary flow regimes. *J. Geophys. Res.,* **93,** 10 927–10 952.

Molteni, F., S. Tibalti, and T. N. Palmer, 1990: Regimes in the wintertime circulation over northern extratropics. I. Observational evidence. *Quart. J. Roy. Meteor. Soc.,* **116,** 31–67.

Monahan, A. H., 2004: A simple model for the skewness of global sea surface winds. *J. Atmos. Sci.,* 61, 2037–2049.

——, L. Pandolfo, and J. C. Fyfe, 2001: The preferred structure of variability of the Northern Hemisphere atmospheric circulation. *Geophys. Res. Lett.,* **28,** 1019–1022.

——, J. C. Fyfe, and L. Pandolfo, 2003: The vertical structure of wintertime climate regimes of the northern hemisphere extratropical atmosphere. *J. Climate,* **16,** 2005–2021.

Neelin, J. D., and W. Weng, 1999: Analytical prototypes for ocean–atmosphere interaction at midlatitudes. Part I. Coupled feedbacks as a sea surface temperature dependent stochastic process. *J. Climate,* **12,** 697–721.

Newman, M., P. D. Sardeshmukh, and C. Penland, 1997: Stochastic forcing of wintertime extratropical flow. *J. Atmos. Sci.,* **54,** 435–455.

——, ——, C. R. Winkler, and J. S. Whitaker, 2003: A study of subseasonal variability. *Mon. Wea. Rev.,* **131,** 1715–1732.

Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.,* **127,** 279–304.

Pandolfo, L., 1993: Observational aspects of the low-frequency intraseasonal variability of the atmosphere in middle latitudes. *Advances in Geophysics,* Vol. **34,** Academic Press, 93–174.

Papanicolaou, G., and W. Kohler, 1974: Asymptotic theory of mixing stochastic differential equations. *Commun. Pure Appl. Math.,* **96,** 641–668.

Park, B. T., and V. Petrosian, 1996: Fokker-Planck equations of stochastic acceleration: A study of numerical methods. *Astrophys. J. Suppl. Ser.,* **103,** 255–267.

Paul, W., and J. Baschnagel, 1999: *Stochastic Processes: From Physics to Finance.* Springer Verlag, 231 pp.

Penland, C., 1996: A stochastic model of Indio Pacific sea surface temperature anomalies. *Physica D,* **98,** 543–558.

——, and L. Matrosova, 1994: A balance condition for stochastic numerical models with application to El Nino–the Southern Oscillation. *J. Climate,* **7,** 1352–1372.

Pérez-Muñuzuri, V., M. N. Lorenzo, P. Montero, K. Fraedrich, E. Kirk, and F. Lunkeit, 2003: Response of a global atmospheric circulation model to spatio-temporal stochastic forcing: Ensemble statistics. *Nonlinear Processes Geophys.,* **10,** 453–461.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran, The Art of Scientific Computing.* Cambridge University Press, 963 pp.

Sardeshmukh, P. D., and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.,* **45,** 1228–1251.

——, G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate,* **13,** 4268–4286.

Sardeshmukh, P., C. Penland, and M. Newman, 2001: Rossby waves in a fluctuating medium. *Stochastic Climate Models,* P. Imkeller and J.-S. von Storch, Eds., Progress in Probability, Birkhäuser Verlag, xx–xx.

Siegert, S., R. Friedrich, and J. Peinke, 1998: Analysis of data sets of stochastic systems. *Phys. Lett. A,* **243,** 275–280.

Smyth, P., K. Ide, and M. Ghil, 1999: Multiple regimes in northern hemisphere height fields via mixture model clustering. *J. Atmos. Sci.,* **56,** 3704–3732.

Sreenivasan, K. R., 1999: Fluid turbulence. *Rev. Mod. Phys.,* **71,** 383–395.

——, and R. A. Antonia, 1997: The phenomenology of small-scale turbulence. *Annu. Rev. Fluid Mech.,* **29,** 435–472.

Stephenson, D. B., A. Hannachi, and A. O'Neill, 2004: On the existence of multiple climate regimes. *Quart. J. Roy. Meteor. Soc.,* **130,** 583–605.

Sura, P., 2002: Noise induced transitions in a barotropic β-plane channel. *J. Atmos. Sci.,* **59,** 97–110.

——, 2003: Stochastic analysis of Southern and Pacific Ocean sea surface winds. *J. Atmos. Sci.,* **60,** 654–666.

——, and J. J. Barsugli, 2002: A note on estimating drift and diffusion parameters from timeseries. *Phys. Lett. A,* **305,** 304–311.

——, and S. T. Gille, 2003: Interpreting wind-driven Southern Ocean variability in a stochastic framework. *J. Mar. Res.,* **61,** 313–334.

Sutera, A., 1986: Probability density distribution of large-scale atmospheric flow. Advances in Geophysics, Vol. XX, Academic Press, 227–249.

van Bebber, W. J., 1891: Die Zugbahnen barometrischer Minima. *Meteor. Z.,* **8,** 361–366.

Weisheimer, A., D. Handorf, and K. Dethloff, 2001: On the structure and variability of atmospheric regimes in coupled climate models. *Atmos. Sci. Lett.,* **2,** 72–80.

Whitaker, J. S., and P. D. Sardeshmukh, 1998: A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci.,* **55,** 237–258.

Wiin-Nielsen, A., 1979: Steady states and stability properties of a low-order barotropic system with forcing and dissipation. *Tellus,* **31,** 375–386.

Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *J. Climate,* **14,** 4474–4493.

Yoden, S., 1985a: Bifurcation properties of a quasi-geostrophic, barotropic, low-order model with topography. *J. Meteor. Soc. Japan,* **63,** 535–546.

——, 1985b: Multiple stable states of quasi-geostrophic barotropic flow over sinusoidal topography. *J. Meteor. Soc. Japan,* **63,** 1031–1044.